CPUE

Alex J. Benecke

February 22, 2018

Contents

Hypothesis

1) H_0: There is no difference in cpe (catch/hour) among years 2013 - 2016.
2) H_0: There is no differenct in cpe for quality length (300mm) and larger largemouth bass among years 2013 - 2016.
3) H_0: There is no difference in cpe for largemouth bass smaller than quality length (300mm) among years 2013 - 2016.

Explaining Analysis

Analysis is divided into three sections. Part 1 uses aov() to test the three hypothesis. Part 2 uses lme() to test the three hypothesis. part 3 uses lme() with year and gcatQ to test the three hypothesis.

I start out simply using aov() to test differences in CPE (Catch/Hour) between years. I sum cpe by site and year and then take the average for each year regardless of size for H_0 1. I run the anova on the yearly average to avoid artificially inflating my sample size (4 years of data instead of 8 to 12 of sites in each year [Unbalanced number of sites]). For H_0 2 and 3 I divide largemouth bass into two categories quality + (greater than or equal to 300mm) and quality - (< 300 mm). Since sites were no largemouth bass were caught cannot influence size structure I removed zeroes from the data. I used aov() as described earlier and tested for differences between years first for Q+ (H_0 2) and then for Q- (H_0 3) largemouth bass.

I also tried to use a lme() model to test the above hypothesis (Part 2) following the above procedure for preparing my data. The only exception is instead of taking a yearly average I sum by site and used sites as a random effect. I log cpe.hr to try and normalize residuals which seems to mostly work (Especially if I remove zeroes [including zeroes and log(cpe.hr +1) skews residuals terribly]).

In part 3 I ran one lme() model with both a year and a gcatQ variable (Factor, 2 levels Q+ and Q-) to test all three hypothesis at once.

Problems and Questions About What I Did

1) Use aov() or lme()?

   I'm unsure if I did the lme() correctly or if it is necessary. The aov() seems overly simple. Am I doing it right?

2) correlation structure for lme()?

   When fitting the lme() in parts 2 and 3 I cannot specify a correlation structure (don't know how or which one to choose). I don't know if this is necessary?

3) pairwise comparison with lme()?

   The pairwise comparison (Year to Year & Q+ Year to Q- Year) doesnt seem to work?

Note: *I am going to remove sites where largemouth bass were not captured because if no larg emouth bass were caught that site can not affect the size structure. I will only do this when comparing gabelhouse length categories.*

Part 1 aov()

1) H_0: There is no difference in cpe (catch/hour) among years 2013 - 2016.

Load and Prepare Data

Load Data

```
cpe <- read.csv("Data/Clean-Data/CPUE_2013-2016.csv") %>% filterD(Species ==
    317)

cpe$Site <- factor(cpe$Site)
```

Sum cpe by Site and Year and Display Data

```
cpeSum <- aggregate(cpe.hr ~ Year + Site, data = cpe, FUN = sum) %>% arrange(Site,
    Year)
```

Year	Site	cpe.hr
2014	1	45.378151
2013	2	86.746988
2014	2	87.032967
2015	2	19.169329
2016	2	59.602649
2013	4	58.009479
2014	4	19.933555
2015	4	31.259045
2016	4	48.949320
2015	5	4.225352
2013	6	4.712042
2014	6	21.021898
2015	6	17.716535
2016	6	44.628099
2013	8	38.709677
2014	8	47.787611
2015	8	117.249698
2016	8	166.591422
2013	10	20.571429
2014	10	27.799228
2016	10	5.872757

	Year	Site	cpe.hr
22	2014	11	117.009751
23	2015	11	33.162743
24	2016	11	26.438188
25	2014	12	36.468886
26	2015	12	12.514484
27	2016	12	10.404624
28	2016	13	0.000000
29	2016	14	8.135593
30	2013	15	81.818182
31	2014	15	40.346409
32	2015	15	25.079164
33	2016	15	16.618581
34	2014	16	45.120859
35	2013	18	97.472924
36	2014	18	16.775396
37	2015	18	64.197531
38	2016	18	148.064516
39	2013	19	0.000000
40	2014	19	0.000000
41	2015	19	0.000000
42	2016	19	0.000000

Average cpe by Year

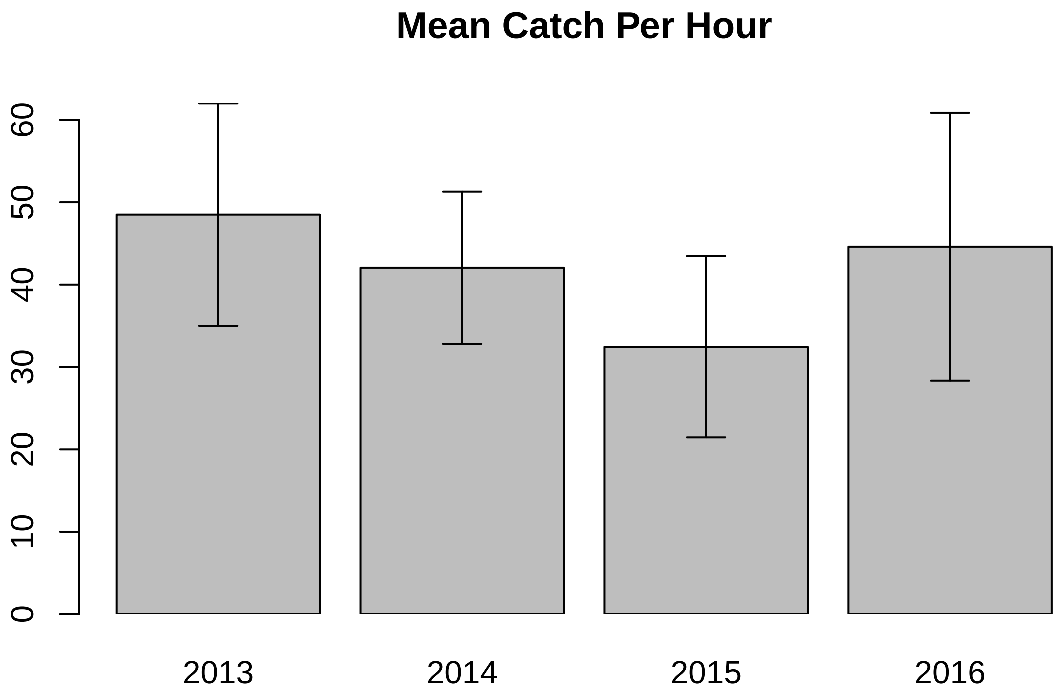
```
(cpeMean <- aggregate(cpe.hr ~ Year, data = cpeSum, FUN = mean))
```

```
##   Year   cpe.hr
## 1 2013 48.50509
## 2 2014 42.05623
## 3 2015 32.45739
## 4 2016 44.60881
```

Test Hypothesis 1

```
aov1 <- aov(cpe.hr ~ Year, data = cpeMean)
summary(aov1)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## Year         1  22.66   22.66   0.386  0.598
## Residuals    2 117.49   58.75
```



Results H_0 1

There is *no significant difference* in CPUE between years ($F_{1,2}$ = 0.386, p = 0.598).

2) H_0: There is no differenct in cpe for quality length (300mm) and larger largemouth bass among years 2013 - 2016.

Load and Prepare Data

Load Data with Gcat and make Q+ and Q-

```
gcat <- read.csv("Data/Clean-Data/CPUE-gcat_2013-2016.csv") %>% filterD(Species ==
    317)

gcat$Site <- factor(gcat$Site)
```

```r
xtabs(caught ~ gcat + Year, data = gcat)
```

```
##            Year
## gcat        2013 2014 2015 2016
##   memorable    0    0    0    0
##   preferred   14   18   15   10
##   quality     39   57   38   47
##   stock       38   65   14   53
##   substock    16    3   13   34
##   trophy       0    0    0    0
```

Make Qcat Variable and Data Frame

```r
Qcat <- gcat %>% mutate(gcatQ = mapvalues(gcat, from = c("substock", "stock",
    "quality", "preferred", "memorable", "trophy"), to = c("quality-", "quality-",
    "quality+", "quality+", "quality+", "quality+"))) %>% dplyr::select(Year,
    Site, gcatQ, cpe.hr)
```

Remove Zeroes

```r
xtabs(cpe.hr ~ Site + Year, data = Qcat)
```

```
##     Year
## Site       2013       2014       2015       2016
##   1    0.000000  45.378151   0.000000   0.000000
##   2   86.746988  87.032967  19.169329  59.602649
##   4   58.009479  19.933555  31.259045  48.949320
##   5    0.000000   0.000000   4.225352   0.000000
##   6    4.712042  21.021898  17.716535  44.628099
##   8   38.709677  47.787611 117.249698 166.591422
##   10  20.571429  27.799228   0.000000   5.872757
##   11   0.000000 117.009751  33.162743  26.438188
##   12   0.000000  36.468886  12.514484  10.404624
##   13   0.000000   0.000000   0.000000   0.000000
##   14   0.000000   0.000000   0.000000   8.135593
##   15  81.818182  40.346409  25.079164  16.618580
##   16   0.000000  45.120859   0.000000   0.000000
##   18  97.472924  16.775396  64.197531 148.064516
##   19   0.000000   0.000000   0.000000   0.000000
```

```r
for(i in 1:length(Qcat$cpe.hr)){
  if(Qcat$cpe.hr[i]==0){
    Qcat$cpe.hr[i] = NA
  } else{
    Qcat$cpe.hr[i] = Qcat$cpe.hr[i]
  }
}

Qcat <- Qcat[!is.na(Qcat$cpe.hr),] ### remove NAs
```

Create Data Frame With Only Quality + Fish

```r
Qpls <- Qcat[Qcat$gcatQ == "quality+", ]
Qpls$gcatQ <- droplevels(Qpls$gcatQ)

str(Qpls)
```

```
## 'data.frame':    53 obs. of  4 variables:
##  $ Year  : int  2013 2013 2013 2013 2013 2013 2013 2013 2013 2014 ...
##  $ Site  : Factor w/ 15 levels "1","2","4","5",..: 2 3 6 7 7 12 12 14 14 1 ...
##  $ gcatQ : Factor w/ 1 level "quality+": 1 1 1 1 1 1 1 1 1 1 ...
##  $ cpe.hr: num  23.66 23.89 8.6 4.11 4.11 ...
```

```r
Qpls.sum <- aggregate(cpe.hr ~ Site + Year, data = Qpls, FUN = sum) %>% arrange(Site,
    Year)
```

Site	Year	cpe.hr
1	2014	11.344538
2	2013	23.658269
2	2014	35.604396
2	2016	23.841060
4	2013	23.886256
4	2014	11.960133
4	2015	26.049204
4	2016	17.799753
5	2015	4.225352
6	2014	15.766423
6	2015	3.543307
6	2016	14.876033
8	2013	8.602151
8	2014	23.893805
8	2015	52.110977
8	2016	24.379233
10	2013	8.228571

	Site	Year	cpe.hr
18	10	2014	22.239382
19	10	2016	5.872757
20	11	2014	39.003250
21	11	2015	29.477994
22	11	2016	17.625459
23	12	2014	31.259045
24	12	2015	12.514484
25	12	2016	10.404624
26	15	2013	53.359684
27	15	2014	22.007132
28	15	2015	25.079164
29	15	2016	4.154645
30	16	2014	38.675022
31	18	2013	68.231047
32	18	2014	10.065238
33	18	2015	59.259259
34	18	2016	84.193548

Average cpe for Q+ by Year

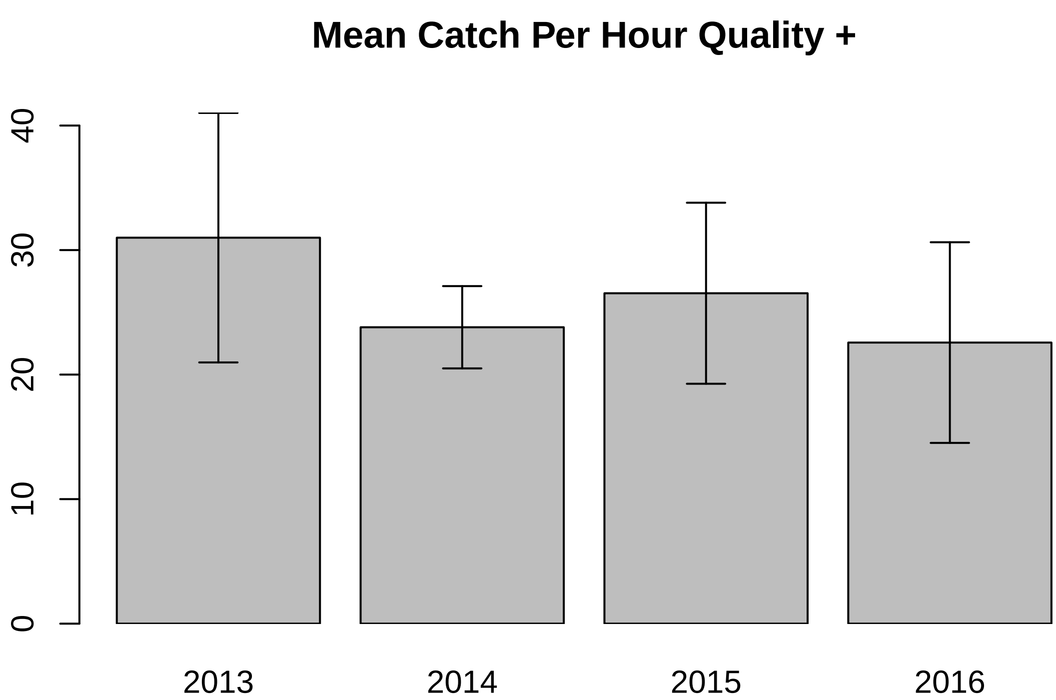
```r
Qpls.mean <- aggregate(cpe.hr ~ Year, data = Qpls.sum, FUN = mean)

xtabs(cpe.hr ~ Year, data = Qpls.mean)
```

```
## Year
##     2013     2014     2015     2016 
## 30.99433 23.80167 26.53247 22.57190
```

Test Hypothesis 2

```r
aov.Qpls <- aov(cpe.hr ~ Year, data = Qpls.mean)
summary(aov.Qpls)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## Year         1  25.39  25.395   3.094  0.221
## Residuals    2  16.41   8.207
```



Results H_0 2

There is *no significant difference* in CPUE for fish > Quality length (300mm) among years 2013 - 2016 ($F_{1,2}$ = 3.09, p = 0.221).

3) H_0: There is no difference in cpe for largemouth bass smaller than quality length (300mm) among years 2013 - 2016.

Load and Prepare Data

Create Q- Data Frame

```
Qless <- Qcat[Qcat$gcatQ == "quality-", ]
Qless$gcatQ <- droplevels(Qless$gcatQ)

str(Qless)
```

```
## 'data.frame':    45 obs. of  4 variables:
##  $ Year  : int  2013 2013 2013 2013 2013 2013 2013 2013 2013 2013 ...
##  $ Site  : Factor w/ 15 levels "1","2","4","5",..: 2 2 3 3 5 6 6 7 12 12 ...
##  $ gcatQ : Factor w/ 1 level "quality-": 1 1 1 1 1 1 1 1 1 1 ...
##  $ cpe.hr: num  35.49 27.6 3.41 30.71 4.71 ...
```

```
Qless.sum <- aggregate(cpe.hr ~ Year + Site, data = Qless, FUN = sum) %>% arrange(Site,
    Year)
```

Year	Site	cpe.hr
2014	1	34.033613
2013	2	63.088718
2014	2	51.428571
2015	2	19.169329
2016	2	35.761589
2013	4	34.123223
2014	4	7.973422
2015	4	5.209841
2016	4	31.149567
2013	6	4.712042
2014	6	5.255475
2015	6	14.173228
2016	6	29.752066
2013	8	30.107527
2014	8	23.893805
2015	8	65.138721

	Year	Site	cpe.hr
17	2016	8	142.212190
18	2013	10	12.342857
19	2014	10	5.559846
20	2014	11	78.006501
21	2015	11	3.684749
22	2016	11	8.812729
23	2014	12	5.209841
24	2016	14	8.135593
25	2013	15	28.458498
26	2014	15	18.339277
27	2016	15	12.463935
28	2014	16	6.445837
29	2013	18	29.241877
30	2014	18	6.710158
31	2015	18	4.938272
32	2016	18	63.870968

Average cpe by Year for Q- fish

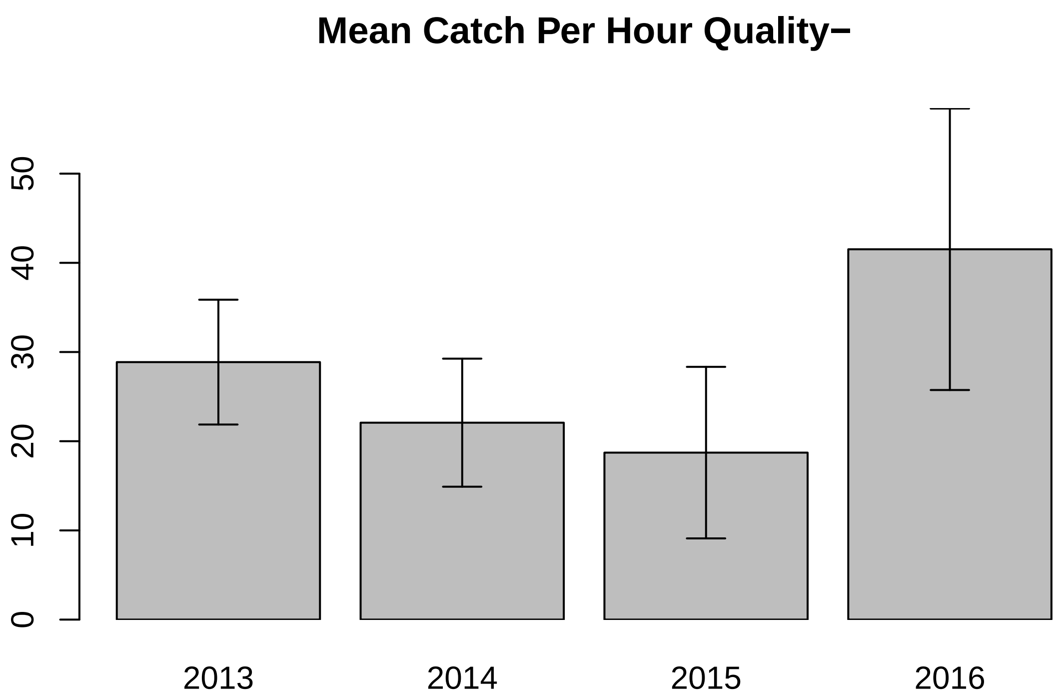
```
Qless.mean <- aggregate(cpe.hr ~ Year, data = Qless.sum, FUN = mean)

xtabs(cpe.hr ~ Year, data = Qless.mean)
```

```
## Year
##     2013     2014     2015     2016
## 28.86782 22.07785 18.71902 41.51983
```

Test Hypothesis 3

```
aov.Qless <- aov(cpe.hr ~ Year, data = Qless.mean)
summary(aov.Qless)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## Year         1  59.85   59.85   0.489  0.557
## Residuals    2 244.73  122.37
```



Results H_0 3

There is *no significat difference* in CPUE for fish < Quality length among years 2013 - 2016 ($F_{1,2}$ = 0.489, p = 0.557).

End of Part 1

Part 2 - Repeated Measures Anova with Mixed effects lme()

Source for following analysis: https://rcompanion.org/handbook/I_09.html

1) H_0: There is no difference in cpe (catch/hour) among years 2013 - 2016.

Load and Prepare Data

Sum CPE by Site and Year without any gabelhouse length categories.

```
cpeSum2 <- aggregate(cpe.hr ~ Site + Year, data = cpe, FUN = sum) %>% arrange(Site,
    Year)
```

Site	Year	cpe.hr
1	2014	45.378151
2	2013	86.746988
2	2014	87.032967
2	2015	19.169329
2	2016	59.602649
4	2013	58.009479
4	2014	19.933555
4	2015	31.259045
4	2016	48.949320
5	2015	4.225352
6	2013	4.712042
6	2014	21.021898
6	2015	17.716535
6	2016	44.628099
8	2013	38.709677
8	2014	47.787611
8	2015	117.249698
8	2016	166.591422
10	2013	20.571429
10	2014	27.799228
10	2016	5.872757

	Site	Year	cpe.hr
22	11	2014	117.009751
23	11	2015	33.162743
24	11	2016	26.438188
25	12	2014	36.468886
26	12	2015	12.514484
27	12	2016	10.404624
28	13	2016	0.000000
29	14	2016	8.135593
30	15	2013	81.818182
31	15	2014	40.346409
32	15	2015	25.079164
33	15	2016	16.618581
34	16	2014	45.120859
35	18	2013	97.472924
36	18	2014	16.775396
37	18	2015	64.197531
38	18	2016	148.064516
39	19	2013	0.000000
40	19	2014	0.000000
41	19	2015	0.000000
42	19	2016	0.000000

```r
cpeSum2$Site <- factor(cpeSum2$Site)
cpeSum2$Year <- factor(cpeSum2$Year)
str(cpeSum2)
```

```
## 'data.frame':    42 obs. of  3 variables:
##  $ Site  : Factor w/ 15 levels "1","2","4","5",..: 1 2 2 2 2 3 3 3 3 4 ...
##  $ Year  : Factor w/ 4 levels "2013","2014",..: 2 1 2 3 4 1 2 3 4 3 ...
##  $ cpe.hr: num  45.4 86.7 87 19.2 59.6 ...
```

Find an initial value for correlation structure.

Note:

I left the code but did not specify a correlation structure in the model.

```r
mod.a = gls(log(cpe.hr+1) ~ Year,
            data = cpeSum2)
ACF(mod.a)

mod.b = lme(log(cpe.hr+1) ~ Year,
            random = ~1|Site,
            data = cpeSum2)
ACF(mod.b)
```

Fit lme() to Test Hypothesis 1

```
# ?corClasses

cpe.mod <- lme(log(cpe.hr + 1) ~ Year, random = ~1 | Site, data = cpeSum2, method = "REML")

Anova(cpe.mod)
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: log(cpe.hr + 1)
##       Chisq Df Pr(>Chisq)
## Year 1.1335  3      0.769
```

```
cpe.fixed <- gls(log(cpe.hr + 1) ~ Year, data = cpeSum2, method = "REML")
Anova(cpe.fixed)
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: log(cpe.hr + 1)
##      Df  Chisq Pr(>Chisq)
## Year  3 0.8743     0.8316
```

```
anova(cpe.mod, cpe.fixed)
```

```
##           Model df      AIC      BIC    logLik   Test  L.Ratio p-value
## cpe.mod       1  6 133.9685 143.7940 -60.98426
## cpe.fixed     2  5 156.2353 164.4232 -73.11763 1 vs 2 24.26674  <.0001
```

Looks like the cpe model with random effects of site is better than the model without random effects (AIC = 133.97, 156.24 respectively). There is *no significant difference* in cpe.hr between years (Analysis of Deviance Table, $X_2 = 1.13$, df = 3, pr(>Chisq) = 0.769).

p-value and pseudo R-squared for model

The nagelkerke function can be used to calculate a p-value and pseudo R-squared value for the model.

```
null.mod <- lme(log(cpe.hr + 1) ~ 1, random = ~1 | Site, data = cpeSum2)
nagelkerke(cpe.mod, null.mod)
```

```
## $Models
##
## Model: "lme.formula, log(cpe.hr + 1) ~ Year, cpeSum2, ~1 | Site, REML"
## Null:  "lme.formula, log(cpe.hr + 1) ~ 1, cpeSum2, ~1 | Site"
##
## $Pseudo.R.squared.for.model.vs.null
##                              Pseudo.R.squared
## McFadden                            0.0103350
## Cox and Snell (ML)                  0.0293455
## Nagelkerke (Cragg and Uhler)        0.0310872
##
## $Likelihood.ratio.test
##  Df.diff LogLik.diff Chisq p.value
##       -3    -0.62548 1.251 0.74081
##
## $Number.of.observations
##
## Model: 42
```

```
## Null:  42
##
## $Messages
## [1] "Note: For models fit with REML, these statistics are based on refitting with ML"
##
## $Warnings
## [1] "None"
```

```r
null.mod2 <- gls(log(cpe.hr + 1) ~ 1, data = cpeSum2)

nagelkerke(cpe.mod, null.mod2)
```

Very poor R^2. Also, not sure which pseudo R^2 to use?

Post-hoc Analysis

```r
leastsquare = lsmeans(cpe.mod,
                      pairwise ~ Year,
                      adjust="tukey")         ###  Tukey-adjusted comparisons

cld(leastsquare,
    alpha   = 0.05,
    Letters = letters,     ### Use lower-case letters for .group
    adjust  = "tukey")     ###  Tukey-adjusted comparisons
```

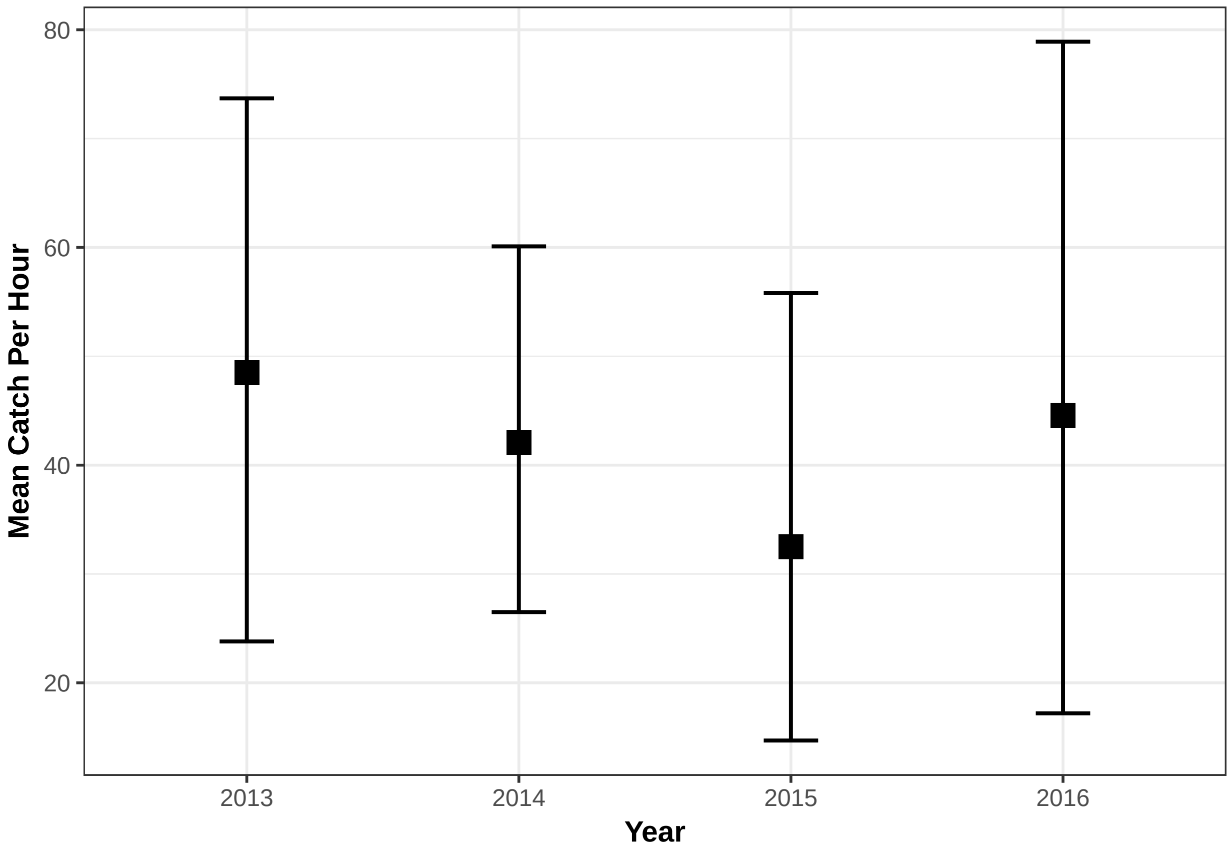
```
##  Year   lsmean        SE df lower.CL upper.CL .group
##  2015 2.759999 0.4124325 14 1.582794 3.937204  a
##  2016 2.918502 0.3966558 14 1.786328 4.050676  a
##  2013 3.056283 0.4330887 14 1.820119 4.292447  a
##  2014 3.078693 0.3966558 14 1.946519 4.210867  a
##
## Results are given on the log (not the response) scale.
## Confidence level used: 0.95
## Conf-level adjustment: sidak method for 4 estimates
## P value adjustment: tukey method for comparing a family of 4 estimates
## significance level used: alpha = 0.05
```

?

Interaction plot

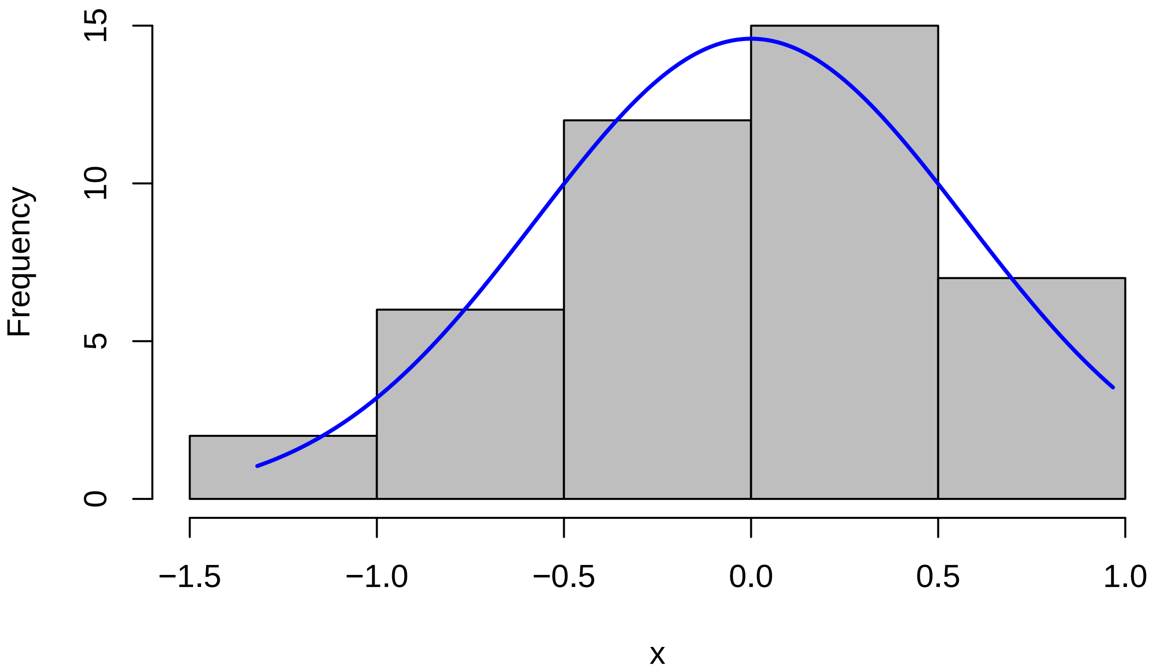
For this plot, we will use the groupwiseMean function to calculate the natural mean of each Instruction x Month combination, along with the confidence interval of each mean with the percentile method.

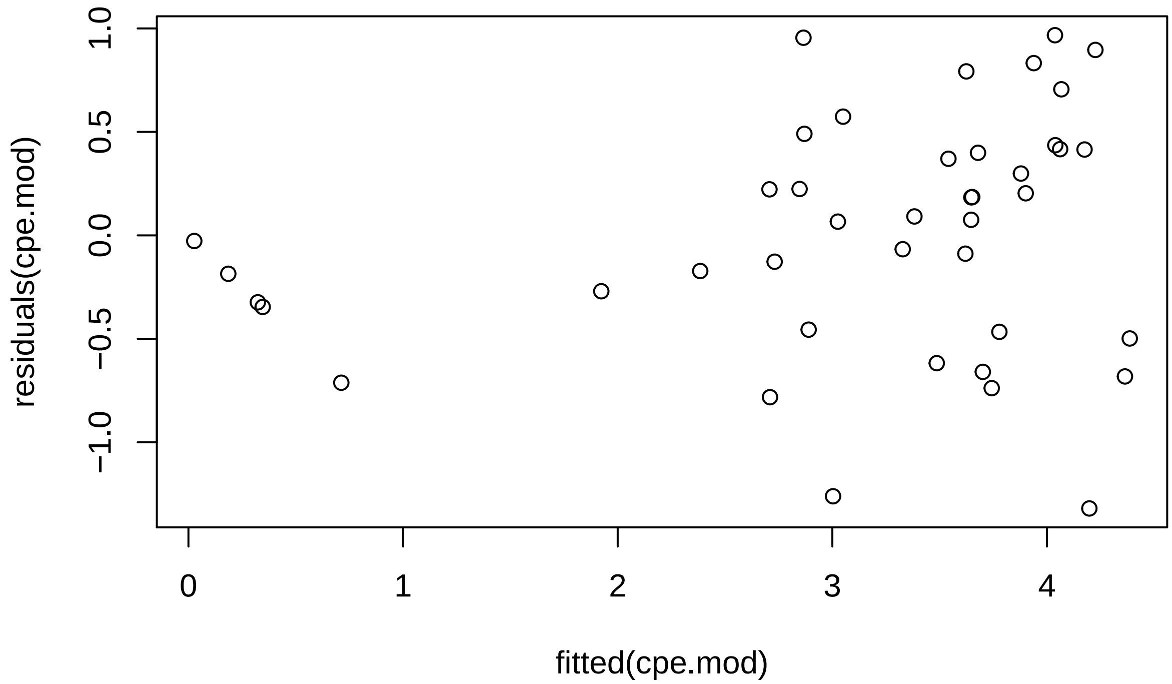
```
##   Year  n Mean Conf.level Percentile.lower Percentile.upper
## 1 2013  8 48.5       0.95             23.8             73.7
## 2 2014 12 42.1       0.95             26.5             60.1
## 3 2015 10 32.5       0.95             14.7             55.8
## 4 2016 12 44.6       0.95             17.2             78.9
```



Histogram of Residuals

Residuals from a mixed model fit with nlme should be normally distributed. Plotting residuals vs. fitted values, to check for homoscedasticity and independence, is probably also advisable.





Results H_0 1

Comming soon...

2) H_0: There is no differenct in cpe for quality length (300mm) and larger largemouth bass among years 2013 - 2016.

Load and Prepare Data

Create Q+ Data Frame and Show Data

```
Qpls.sum %>% arrange(Site, Year)
```

Site	Year	cpe.hr
1	2014	11.344538
2	2013	23.658269
2	2014	35.604396
2	2016	23.841060
4	2013	23.886256
4	2014	11.960133
4	2015	26.049204
4	2016	17.799753
5	2015	4.225352
6	2014	15.766423
6	2015	3.543307
6	2016	14.876033
8	2013	8.602151
8	2014	23.893805
8	2015	52.110977
8	2016	24.379233
10	2013	8.228571

	Site	Year	cpe.hr
18	10	2014	22.239382
19	10	2016	5.872757
20	11	2014	39.003250
21	11	2015	29.477994
22	11	2016	17.625459
23	12	2014	31.259045
24	12	2015	12.514484
25	12	2016	10.404624
26	15	2013	53.359684
27	15	2014	22.007132
28	15	2015	25.079164
29	15	2016	4.154645
30	16	2014	38.675022
31	18	2013	68.231047
32	18	2014	10.065238
33	18	2015	59.259259
34	18	2016	84.193548

```r
Qpls.sum$Site <- factor(Qpls.sum$Site)
Qpls.sum$Year <- factor(Qpls.sum$Year)
str(Qpls.sum)
```

```
## 'data.frame':    34 obs. of  3 variables:
##  $ Site  : Factor w/ 12 levels "1","2","4","5",..: 1 2 2 2 3 3 3 3 4 5 ...
##  $ Year  : Factor w/ 4 levels "2013","2014",..: 2 1 2 4 1 2 3 4 3 2 ...
##  $ cpe.hr: num  11.3 23.7 35.6 23.8 23.9 ...
```

Value for Correlation Structure (Not Used)

```r
mod.a.Qpls = gls(log(cpe.hr) ~ Year, data = Qpls.sum)
ACF(mod.a.Qpls)

mod.b.Qpls = lme(log(cpe.hr) ~ Year, random = ~1 | Site, data = Qpls.sum)
ACF(mod.b.Qpls)
```

Fit lme() to Test Hypothesis 2

```r
Qpls.mod <- lme(log(cpe.hr) ~ Year, random = ~1 | Site, data = Qpls.sum, method = "REML")

Anova(Qpls.mod)
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: log(cpe.hr)
##       Chisq Df Pr(>Chisq)
## Year 1.0627  3     0.7861
```

```r
Qpls.fixed <- gls(log(cpe.hr) ~ Year, data = Qpls.sum, method = "REML")
Anova(Qpls.fixed)
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: log(cpe.hr)
##      Df  Chisq Pr(>Chisq)
## Year  3 1.0414     0.7912
```

```r
anova(Qpls.mod, Qpls.fixed)
```

```
##            Model df      AIC       BIC    logLik   Test   L.Ratio p-value
## Qpls.mod       1  6 93.95703 102.36421 -40.97851
## Qpls.fixed     2  5 92.15914  99.16513 -41.07957 1 vs 2 0.2021133   0.653
```

Pseudo R-squared

```
null.Qpls.mod <- lme(log(cpe.hr) ~ 1, random = ~1 | Site, data = Qpls.sum)
nagelkerke(Qpls.mod, null.Qpls.mod)
```

```
## $Models
##
## Model: "lme.formula, log(cpe.hr) ~ Year, Qpls.sum, ~1 | Site, REML"
## Null:  "lme.formula, log(cpe.hr) ~ 1, Qpls.sum, ~1 | Site"
##
## $Pseudo.R.squared.for.model.vs.null
##                              Pseudo.R.squared
## McFadden                            0.0152657
## Cox and Snell (ML)                  0.0353903
## Nagelkerke (Cragg and Uhler)        0.0390790
##
## $Likelihood.ratio.test
##  Df.diff LogLik.diff  Chisq p.value
##       -3    -0.61254 1.2251   0.747
##
## $Number.of.observations
##
## Model: 34
## Null:  34
##
## $Messages
## [1] "Note: For models fit with REML, these statistics are based on refitting with ML"
##
## $Warnings
## [1] "None"
```

```
null.Qpls.fixed <- gls(log(cpe.hr) ~ 1, data = Qpls.sum)

nagelkerke(Qpls.mod, null.Qpls.fixed)
```

```
## $Models
##
## Model: "lme.formula, log(cpe.hr) ~ Year, Qpls.sum, ~1 | Site, REML"
## Null:  "gls, log(cpe.hr) ~ 1, Qpls.sum"
##
## $Pseudo.R.squared.for.model.vs.null
##                              Pseudo.R.squared
## McFadden                            0.0173817
## Cox and Snell (ML)                  0.0402807
## Nagelkerke (Cragg and Uhler)        0.0444557
##
## $Likelihood.ratio.test
##  Df.diff LogLik.diff  Chisq p.value
##       -4    -0.69895 1.3979 0.84456
##
## $Number.of.observations
##
## Model: 34
## Null:  34
##
```

```
## $Messages
## [1] "Note: For models fit with REML, these statistics are based on refitting with ML"
##
## $Warnings
## [1] "None"
```

Post-hoc analysis

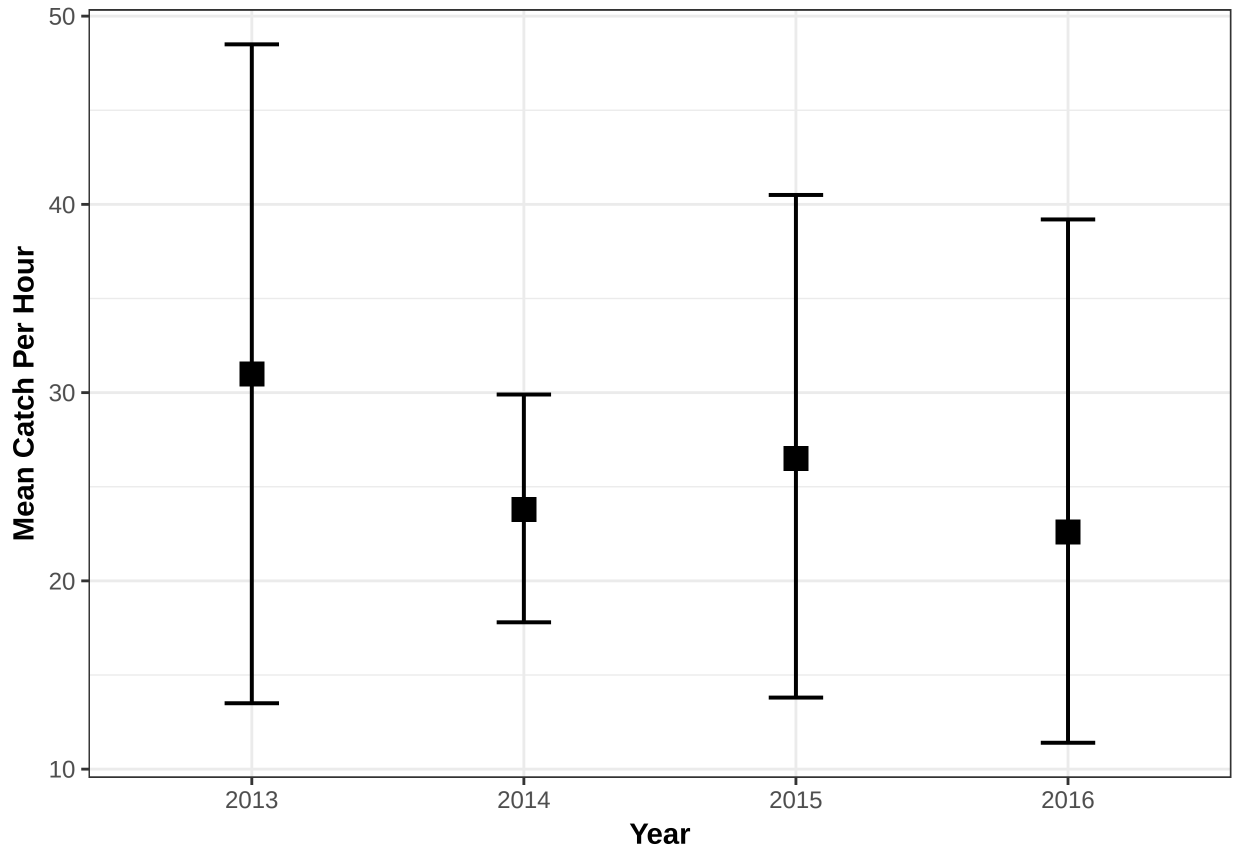
```
leastsquare.Qpls.mod = lsmeans(Qpls.mod,
                      pairwise ~ Year,
                      adjust="tukey")         ###  Tukey-adjusted comparisons

cld(leastsquare.Qpls.mod,
    alpha   = 0.05,
    Letters = letters,     ### Use lower-case letters for .group
    adjust  = "tukey")     ###  Tukey-adjusted comparisons
```

```
##  Year   lsmean        SE df lower.CL upper.CL .group
##  2016 2.739642 0.2753030 11 1.921945 3.557339  a
##  2015 2.883354 0.2916444 11 2.017120 3.749587  a
##  2014 3.050238 0.2492053 11 2.310056 3.790421  a
##  2013 3.098501 0.3362694 11 2.099723 4.097278  a
##
## Results are given on the log (not the response) scale.
## Confidence level used: 0.95
## Conf-level adjustment: sidak method for 4 estimates
## P value adjustment: tukey method for comparing a family of 4 estimates
## significance level used: alpha = 0.05
```

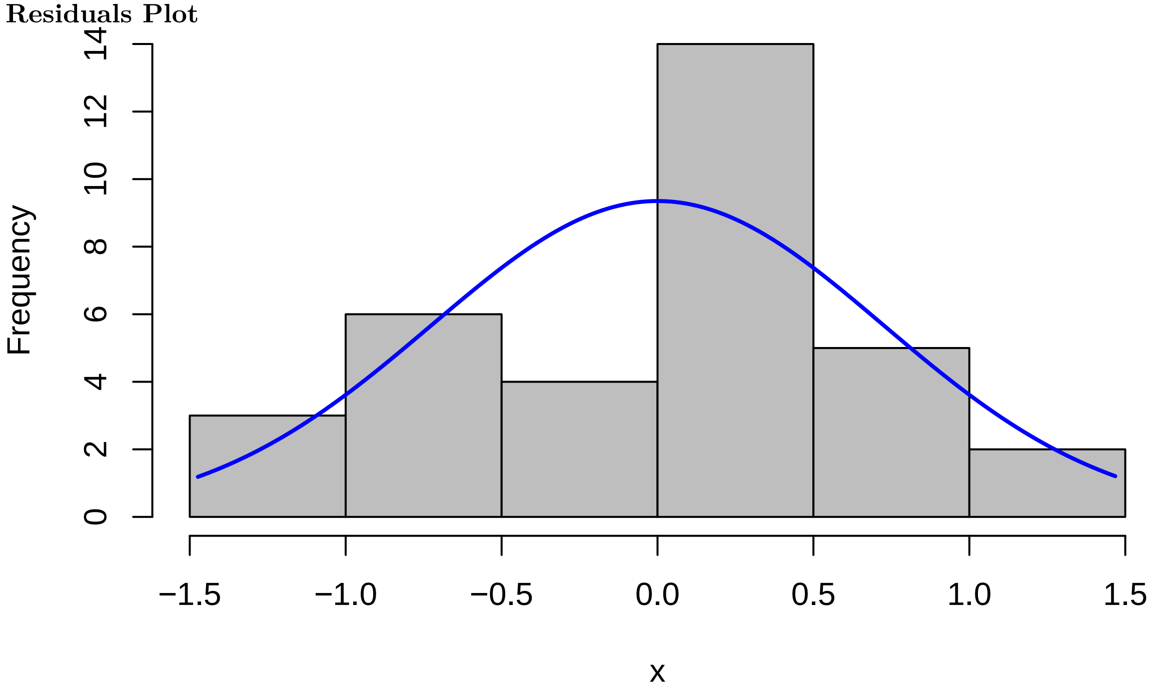
This isn't working right ^

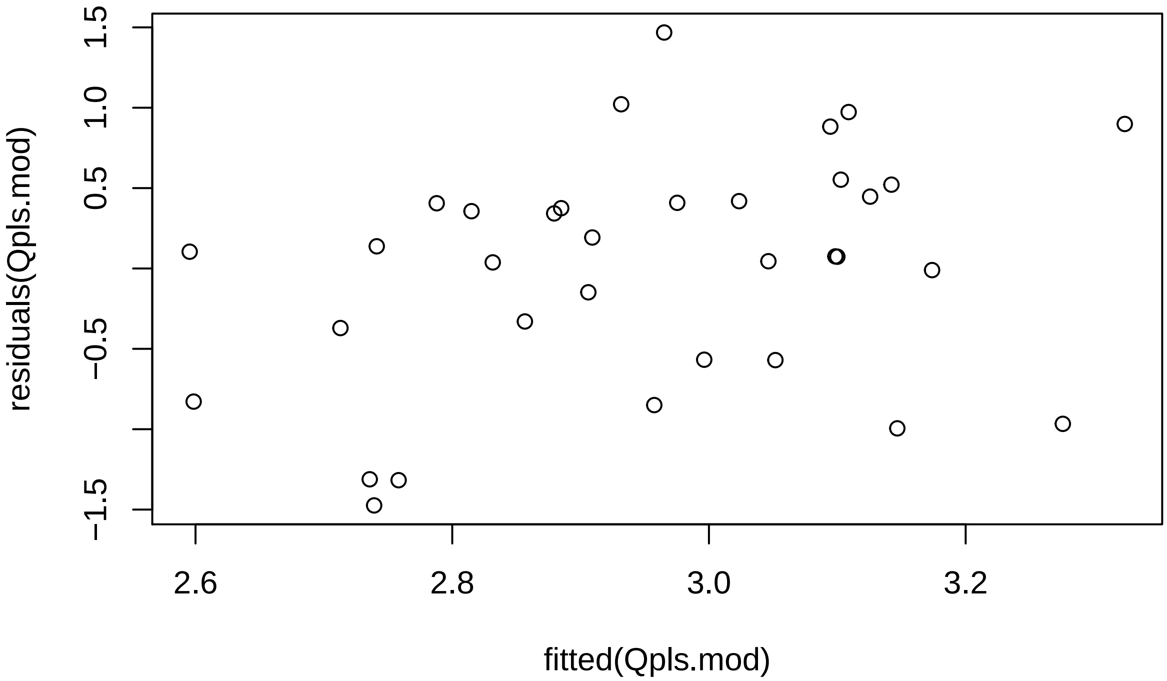
Interaction plot

```
##   Year  n Mean Conf.level Percentile.lower Percentile.upper
## 1 2013  6 31.0       0.95             13.5             48.5
## 2 2014 11 23.8       0.95             17.8             29.9
## 3 2015  8 26.5       0.95             13.8             40.5
## 4 2016  9 22.6       0.95             11.4             39.2
```



Residuals Plot





Pretty normal.

Results H_0 2

Comming soon...

3) H_0: There is no difference in cpe for largemouth bass smaller than quality length (300mm) among years 2013 - 2016.

Load and Prepare Data

Create Q- data Frame and View

```
Qless.sum %>% arrange(Site, Year)
```

Year	Site	cpe.hr
2014	1	34.033613
2013	2	63.088718
2014	2	51.428571
2015	2	19.169329
2016	2	35.761589
2013	4	34.123223
2014	4	7.973422
2015	4	5.209841
2016	4	31.149567
2013	6	4.712042
2014	6	5.255475
2015	6	14.173228
2016	6	29.752066
2013	8	30.107527
2014	8	23.893805
2015	8	65.138721

	Year	Site	cpe.hr
17	2016	8	142.212190
18	2013	10	12.342857
19	2014	10	5.559846
20	2014	11	78.006501
21	2015	11	3.684749
22	2016	11	8.812729
23	2014	12	5.209841
24	2016	14	8.135593
25	2013	15	28.458498
26	2014	15	18.339277
27	2016	15	12.463935
28	2014	16	6.445837
29	2013	18	29.241877
30	2014	18	6.710158
31	2015	18	4.938272
32	2016	18	63.870968

```
Qless.sum$Site <- factor(Qless.sum$Site)
Qless.sum$Year <- factor(Qless.sum$Year)
str(Qless.sum)
```

```
## 'data.frame':    32 obs. of  3 variables:
##  $ Year  : Factor w/ 4 levels "2013","2014",..: 2 1 2 3 4 1 2 3 4 1 ...
##  $ Site  : Factor w/ 12 levels "1","2","4","6",..: 1 2 2 2 2 3 3 3 3 4 ...
##  $ cpe.hr: num  34 63.1 51.4 19.2 35.8 ...
```

Value for Correlation Structure (Not Used)

```
mod.a.Qlss = gls(log(cpe.hr) ~ Year, data = Qless.sum)
ACF(mod.a.Qlss)

mod.b.Qlss = lme(log(cpe.hr) ~ Year, random = ~1 | Site, data = Qless.sum)
ACF(mod.b.Qlss)
```

Fit lme() to Test Hypothesis 3

```r
# ?corClasses

Qlss.mod <- lme(log(cpe.hr) ~ Year, random = ~1 | Site, data = Qless.sum, method = "REML")

Anova(Qlss.mod)
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: log(cpe.hr)
##      Chisq Df Pr(>Chisq)
## Year 5.179  3     0.1591
```

```r
Qlss.fixed <- gls(log(cpe.hr) ~ Year, data = Qless.sum, method = "REML")
Anova(Qlss.fixed)
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: log(cpe.hr)
##      Df  Chisq Pr(>Chisq)
## Year  3 4.0789     0.2531
```

```r
anova(Qlss.mod, Qlss.fixed)
```

```
##            Model df      AIC      BIC    logLik   Test  L.Ratio p-value
## Qlss.mod       1  6 97.62422 105.6175 -42.81211
## Qlss.fixed     2  5 96.77762 103.4386 -43.38881 1 vs 2 1.153399  0.2828
```

Pseudo R-Squared

```r
null.Qlss.mod <- lme(log(cpe.hr) ~ 1, random = ~1 | Site, data = Qless.sum)
nagelkerke(Qlss.mod, null.Qlss.mod)
```

```
## $Models
##
## Model: "lme.formula, log(cpe.hr) ~ Year, Qless.sum, ~1 | Site, REML"
## Null:  "lme.formula, log(cpe.hr) ~ 1, Qless.sum, ~1 | Site"
##
## $Pseudo.R.squared.for.model.vs.null
##                              Pseudo.R.squared
## McFadden                            0.0588341
## Cox and Snell (ML)                  0.1515870
## Nagelkerke (Cragg and Uhler)        0.1614640
##
## $Likelihood.ratio.test
##  Df.diff LogLik.diff  Chisq p.value
##       -3     -2.6302 5.2604 0.15369
##
## $Number.of.observations
##
## Model: 32
## Null:  32
##
## $Messages
## [1] "Note: For models fit with REML, these statistics are based on refitting with ML"
##
## $Warnings
```

```
## [1] "None"
```

```r
null.Qlss.fixed <- gls(log(cpe.hr) ~ 1, data = Qless.sum)

nagelkerke(Qlss.mod, null.Qlss.fixed)
```

```
## $Models
##
## Model: "lme.formula, log(cpe.hr) ~ Year, Qless.sum, ~1 | Site, REML"
## Null:  "gls, log(cpe.hr) ~ 1, Qless.sum"
##
## $Pseudo.R.squared.for.model.vs.null
##                              Pseudo.R.squared
## McFadden                             0.063585
## Cox and Snell (ML)                   0.163529
## Nagelkerke (Cragg and Uhler)         0.174025
##
## $Likelihood.ratio.test
##  Df.diff LogLik.diff Chisq p.value
##       -4      -2.857 5.714 0.22155
##
## $Number.of.observations
##
## Model: 32
## Null:  32
##
## $Messages
## [1] "Note: For models fit with REML, these statistics are based on refitting with ML"
##
## $Warnings
## [1] "None"
```

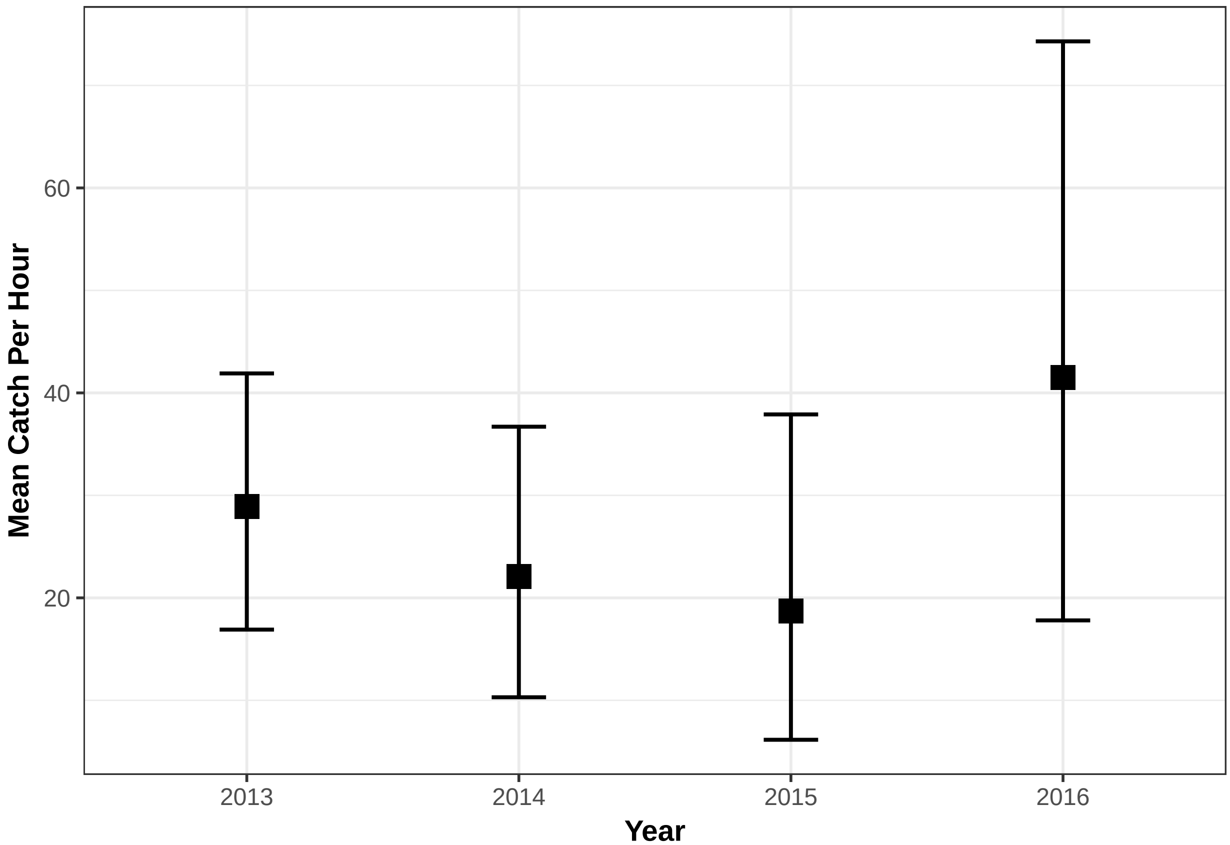
Post-hoc analysis

```r
leastsquare.Qlss.mod = lsmeans(Qlss.mod,
                      pairwise ~ Year,
                      adjust="tukey")         ###  Tukey-adjusted comparisons

cld(leastsquare.Qlss.mod,
    alpha   = 0.05,
    Letters = letters,     ### Use lower-case letters for .group
    adjust  = "tukey")     ###  Tukey-adjusted comparisons
```

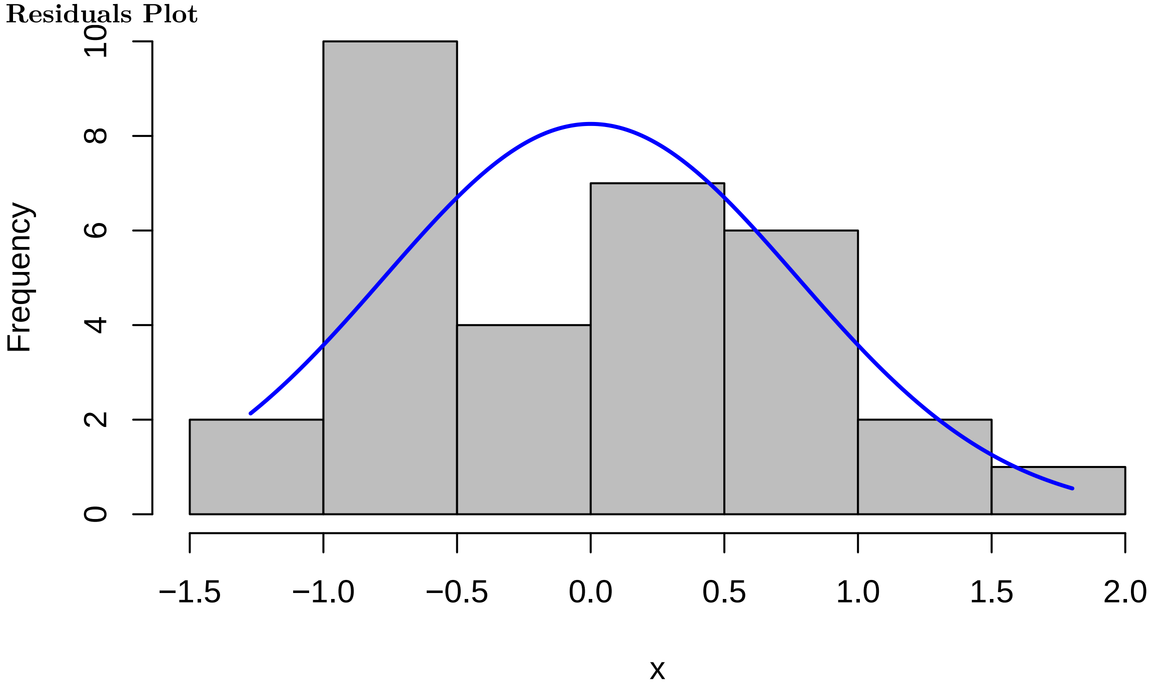
```
##  Year   lsmean        SE df lower.CL upper.CL .group
##  2015 2.283130 0.3948605 11 1.110327 3.455933  a
##  2014 2.595112 0.2957715 11 1.716621 3.473604  a
##  2013 3.063977 0.3669811 11 1.973981 4.153974  a
##  2016 3.241295 0.3446474 11 2.217633 4.264957  a
##
## Results are given on the log (not the response) scale.
## Confidence level used: 0.95
## Conf-level adjustment: sidak method for 4 estimates
## P value adjustment: tukey method for comparing a family of 4 estimates
## significance level used: alpha = 0.05
```

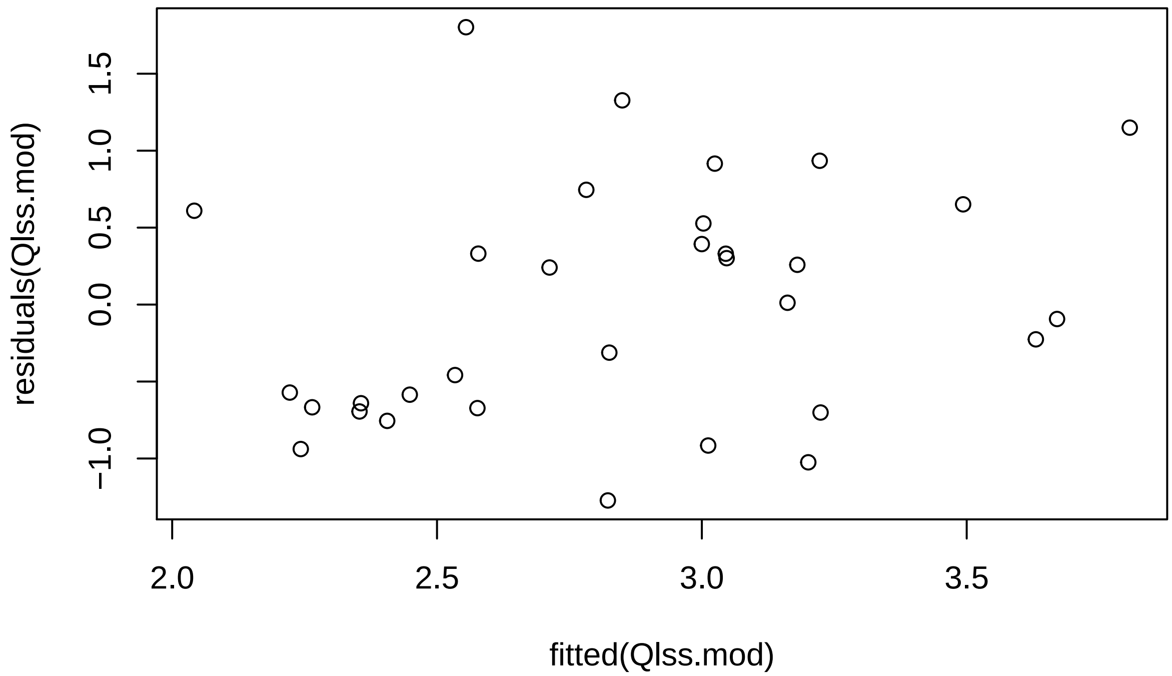
Interaction plot

```
##   Year  n Mean Conf.level Percentile.lower Percentile.upper
## 1 2013  7 28.9       0.95            16.90             41.9
## 2 2014 11 22.1       0.95            10.30             36.7
## 3 2015  6 18.7       0.95             6.15             37.9
## 4 2016  8 41.5       0.95            17.80             74.3
```



Residuals Plot





Almost normal

Results H_0 3

Comming soon…

End of Part 2

Part 3 - Repeated Measures ANOVA lme(cpe.hr ~ Year + GcatQ + Year*GcatQ)

Load and Prepare Data

Create QcatSum Object and View Data Find the sum of largemouth bass by gcatQ (Q+ and Q-) by site for each Year.

```
QcatSum <- aggregate(cpe.hr ~ Year + Site + gcatQ, data = Qcat, FUN = sum)
```

Year	Site	gcatQ	cpe.hr
2014	1	quality+	11.344538
2013	2	quality+	23.658269
2014	2	quality+	35.604396
2016	2	quality+	23.841060
2013	4	quality+	23.886256
2014	4	quality+	11.960133
2015	4	quality+	26.049204
2016	4	quality+	17.799753
2015	5	quality+	4.225352
2014	6	quality+	15.766423
2015	6	quality+	3.543307
2016	6	quality+	14.876033
2013	8	quality+	8.602151
2014	8	quality+	23.893805
2015	8	quality+	52.110977
2016	8	quality+	24.379233
2013	10	quality+	8.228571
2014	10	quality+	22.239382
2016	10	quality+	5.872757
2014	11	quality+	39.003250
2015	11	quality+	29.477994
2016	11	quality+	17.625459
2014	12	quality+	31.259045
2015	12	quality+	12.514484
2016	12	quality+	10.404624
2013	15	quality+	53.359684
2014	15	quality+	22.007132
2015	15	quality+	25.079164
2016	15	quality+	4.154645
2014	16	quality+	38.675022
2013	18	quality+	68.231047
2014	18	quality+	10.065238
2015	18	quality+	59.259259

	Year	Site	gcatQ	cpe.hr
34	2016	18	quality+	84.193548
35	2014	1	quality-	34.033613
36	2013	2	quality-	63.088718
37	2014	2	quality-	51.428571
38	2015	2	quality-	19.169329
39	2016	2	quality-	35.761589
40	2013	4	quality-	34.123223
41	2014	4	quality-	7.973422
42	2015	4	quality-	5.209841
43	2016	4	quality-	31.149567
44	2013	6	quality-	4.712042
45	2014	6	quality-	5.255475
46	2015	6	quality-	14.173228
47	2016	6	quality-	29.752066
48	2013	8	quality-	30.107527
49	2014	8	quality-	23.893805
50	2015	8	quality-	65.138721
51	2016	8	quality-	142.212190
52	2013	10	quality-	12.342857
53	2014	10	quality-	5.559846
54	2014	11	quality-	78.006501
55	2015	11	quality-	3.684749
56	2016	11	quality-	8.812729
57	2014	12	quality-	5.209841
58	2016	14	quality-	8.135593
59	2013	15	quality-	28.458498
60	2014	15	quality-	18.339277
61	2016	15	quality-	12.463935
62	2014	16	quality-	6.445837
63	2013	18	quality-	29.241877
64	2014	18	quality-	6.710158
65	2015	18	quality-	4.938272
66	2016	18	quality-	63.870968

```r
QcatSum$Year <- factor(QcatSum$Year)
str(QcatSum)
```

```
## 'data.frame':    66 obs. of  4 variables:
##  $ Year  : Factor w/ 4 levels "2013","2014",..: 2 1 2 4 1 2 3 4 3 2 ...
##  $ Site  : Factor w/ 15 levels "1","2","4","5",..: 1 2 2 2 3 3 3 3 4 5 ...
##  $ gcatQ : Factor w/ 2 levels "quality+","quality-": 1 1 1 1 1 1 1 1 1 1 ...
##  $ cpe.hr: num  11.3 23.7 35.6 23.8 23.9 ...
```

Value for Correlation Structure (Not Used)

```r
mod.a.Q = gls(log(cpe.hr) ~ Year, data = QcatSum)
ACF(mod.a.Q)

mod.b.Q = lme(log(cpe.hr) ~ Year, random = ~1 | Site, data = QcatSum)
ACF(mod.b.Q)
```

Fit lme() to Test All Three Hypothesis

```
Q.mod <- lme(log(cpe.hr) ~ gcatQ + Year + gcatQ/Year, random = ~1 | Site, data = QcatSum,
    method = "REML")

Anova(Q.mod)
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: log(cpe.hr)
##             Chisq Df Pr(>Chisq)
## gcatQ      0.4791  1     0.4888
## Year       2.1943  3     0.5331
## gcatQ:Year 4.5106  3     0.2113
```

```
Q.fixed <- gls(log(cpe.hr) ~ gcatQ + Year + gcatQ * Year, data = QcatSum, method = "REML")
Anova(Q.fixed)
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: log(cpe.hr)
##            Df  Chisq Pr(>Chisq)
## gcatQ       1 0.2780     0.5980
## Year        3 2.0867     0.5546
## gcatQ:Year  3 3.5927     0.3089
```

```
anova(Q.mod, Q.fixed)
```

```
##         Model df      AIC      BIC    logLik   Test  L.Ratio p-value
## Q.mod       1 10 187.2468 207.8512 -83.62339
## Q.fixed     2  9 187.8197 206.3637 -84.90988 1 vs 2 2.572966  0.1087
```

Pseudo R-Squared

```
null.Q.mod <- lme(log(cpe.hr) ~ 1, random = ~1 | Site, data = QcatSum)
nagelkerke(Q.mod, null.Q.mod)
```

```
## $Models
##
## Model: "lme.formula, log(cpe.hr) ~ gcatQ + Year + gcatQ/Year, QcatSum, ~1 | Site, REML"
## Null:  "lme.formula, log(cpe.hr) ~ 1, QcatSum, ~1 | Site"
##
## $Pseudo.R.squared.for.model.vs.null
##                              Pseudo.R.squared
## McFadden                            0.0441903
## Cox and Snell (ML)                  0.1077340
## Nagelkerke (Cragg and Uhler)        0.1165710
##
## $Likelihood.ratio.test
##  Df.diff LogLik.diff  Chisq p.value
##       -7     -3.7617 7.5234 0.37649
##
## $Number.of.observations
##
## Model: 66
## Null:  66
##
## $Messages
```

```
## [1] "Note: For models fit with REML, these statistics are based on refitting with ML"
##
## $Warnings
## [1] "None"
```

```
null.Q.fixed <- gls(log(cpe.hr) ~ 1, data = QcatSum)

nagelkerke(Q.mod, null.Q.fixed)
```

```
## $Models
##
## Model: "lme.formula, log(cpe.hr) ~ gcatQ + Year + gcatQ/Year, QcatSum, ~1 | Site, REML"
## Null:  "gls, log(cpe.hr) ~ 1, QcatSum"
##
## $Pseudo.R.squared.for.model.vs.null
##                              Pseudo.R.squared
## McFadden                            0.0544962
## Cox and Snell (ML)                  0.1324720
## Nagelkerke (Cragg and Uhler)        0.1430130
##
## $Likelihood.ratio.test
##  Df.diff LogLik.diff  Chisq p.value
##       -8     -4.6895 9.3791 0.31133
##
## $Number.of.observations
##
## Model: 66
## Null:  66
##
## $Messages
## [1] "Note: For models fit with REML, these statistics are based on refitting with ML"
##
## $Warnings
## [1] "None"
```

Post-hoc analysis

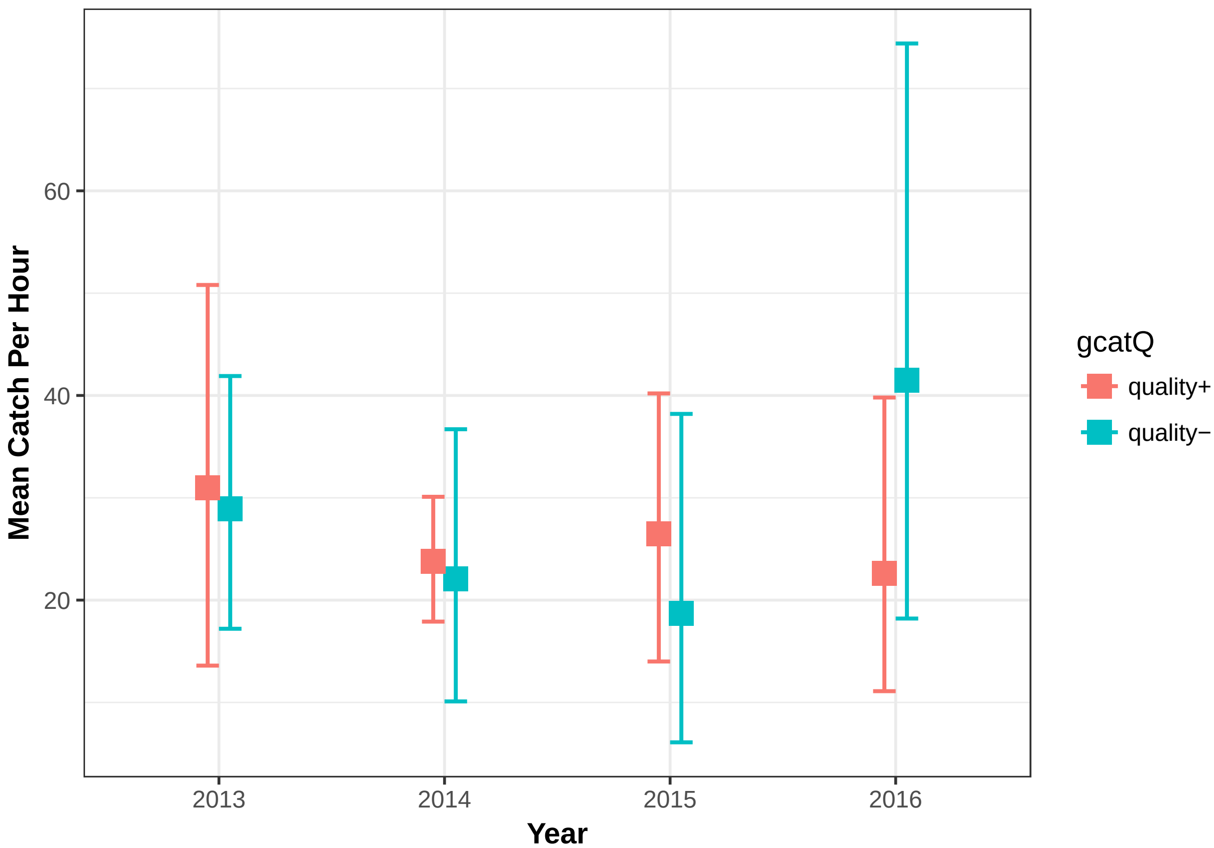
```
leastsquare.Q.mod = lsmeans(Q.mod,
                      pairwise ~ gcatQ:Year,
                      adjust="tukey")         ###  Tukey-adjusted comparisons

cld(leastsquare.Q.mod,
    alpha   = 0.05,
    Letters = letters,     ### Use lower-case letters for .group
    adjust  = "tukey")     ###  Tukey-adjusted comparisons
```

```
##  gcatQ    Year   lsmean        SE df lower.CL upper.CL .group
##  quality- 2015 2.254721 0.3635786 12 1.056466 3.452977  a
##  quality- 2014 2.579502 0.2727896 12 1.680462 3.478542  a
##  quality+ 2016 2.713143 0.3003457 12 1.723286 3.703001  a
##  quality+ 2015 2.882990 0.3168335 12 1.838793 3.927187  a
##  quality+ 2013 3.015239 0.3636609 12 1.816712 4.213765  a
##  quality+ 2014 3.025156 0.2727896 12 2.126116 3.924196  a
##  quality- 2013 3.066420 0.3381391 12 1.952006 4.180834  a
##  quality- 2016 3.211019 0.3171042 12 2.165931 4.256108  a
##
```

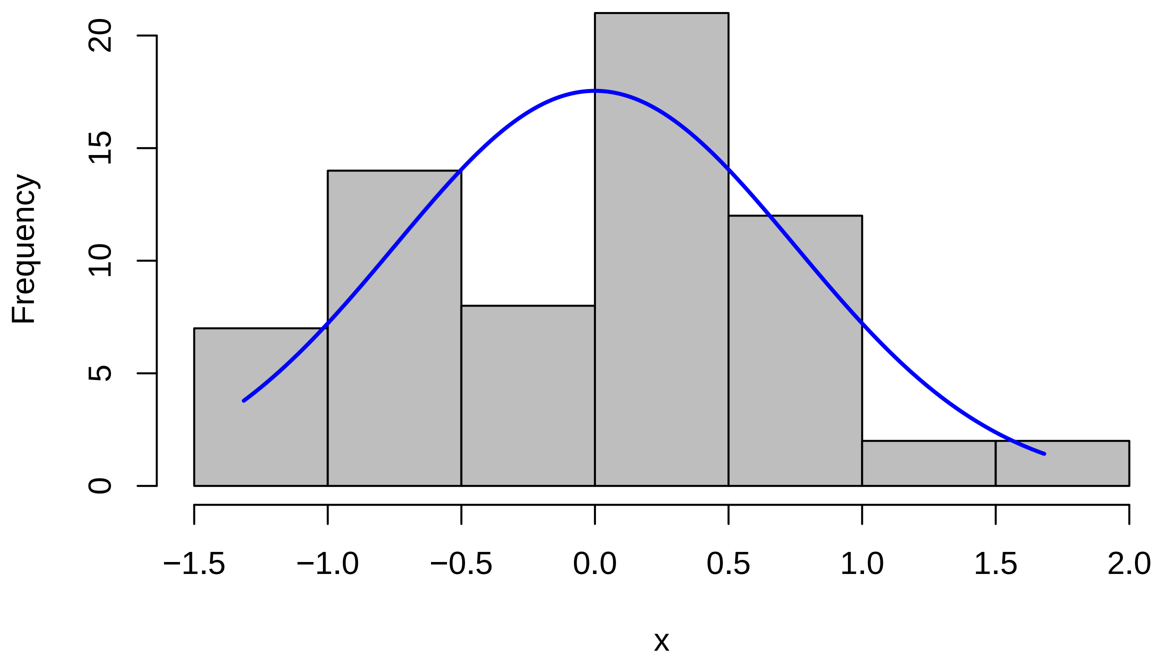
```
## Results are given on the log (not the response) scale.
## Confidence level used: 0.95
## Conf-level adjustment: sidak method for 8 estimates
## P value adjustment: tukey method for comparing a family of 8 estimates
## significance level used: alpha = 0.05
```

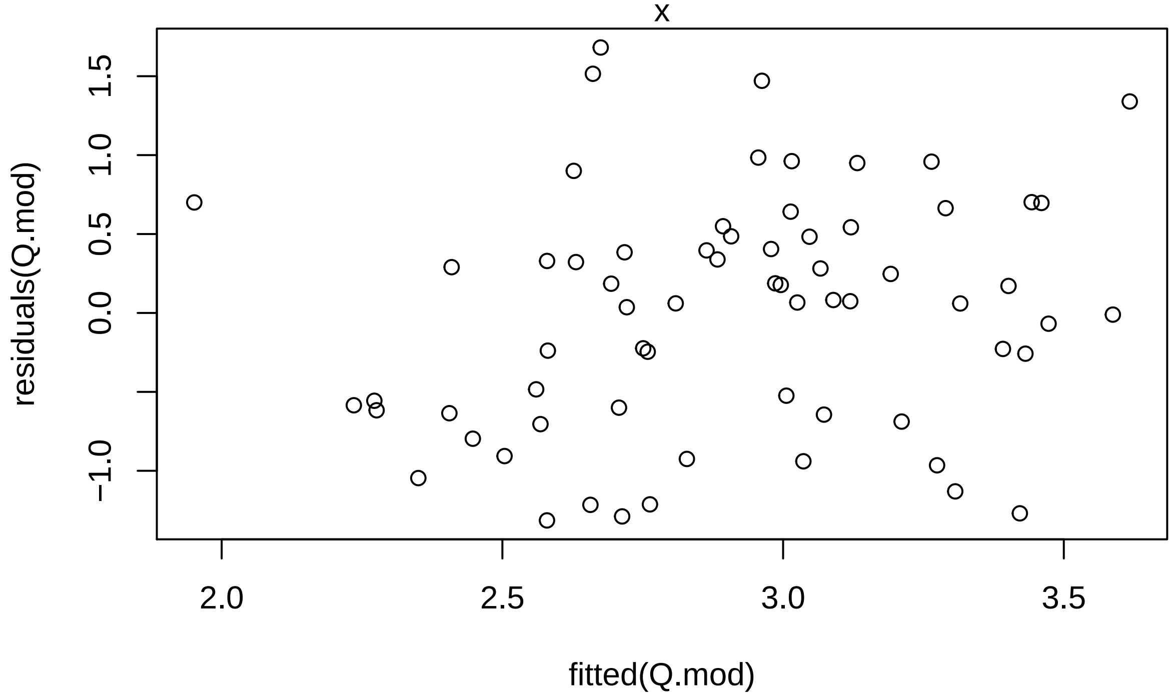
Interaction plot

```
##      gcatQ Year  n Mean Conf.level Percentile.lower Percentile.upper
## 1 quality+ 2013  6 31.0       0.95             13.6             50.8
## 2 quality+ 2014 11 23.8       0.95             17.9             30.1
## 3 quality+ 2015  8 26.5       0.95             14.0             40.2
## 4 quality+ 2016  9 22.6       0.95             11.1             39.8
## 5 quality- 2013  7 28.9       0.95             17.2             41.9
## 6 quality- 2014 11 22.1       0.95             10.1             36.7
## 7 quality- 2015  6 18.7       0.95              6.1             38.2
## 8 quality- 2016  8 41.5       0.95             18.2             74.4
```



Residuals Plot





sorta normal

Results Part 3 All Hypothesis